International Journal of Advance Research in Science and Engineering
Vol. No.4, Special Issue (01), August 2015
www.ijarse.com
IJARSE
ISSN 2319 - 8354

A SURVEY ON BRAIN TUMOR MEDICAL IMAGE CLASSIFICATION

Nayana Suresh

Computer Science and Engineering, Vimal Jyothi Engineering College, Kannur, Kerala, (India)

ABSTRACT

Brain tumors are the most aggressive and devastating types of cancer and therefore, its correct identification at anearly stage followed by treatment is its only cure. Brain is the central part of the body which is responsible for controlling and coordinating all other body organs, so if a tumor is located in a portion of brain then activities which are controlled by that part of nervous system are also affected. Tumor has a variant and complex structure and hence its classification is difficult. The classification of brain tumors consists of several steps including segmentation, feature extraction, and classification model construction. The objective of this survey is to find out the algorithms and approaches that have been used for the classification of brain tumors in Computed Tomography or Magnetic Resonance images, and analyze already proposed systems and will try to find the efficient and effective approaches.

I. INTRODUCTION

Medical image analysis[2] can be used as preliminary screening techniques to help doctors. Various aspects of segmentation features and algorithms have been extensively explored for many years in a host of publications. However, the problem remains challenging, with no general and unique solution, due to a large and constantly growing number of different objects of interest, large variations of their properties in images, different medical imaging modalities, and associated changes of signal homogeneity, variability, and noise for each object. Computed Tomography(CT) and Magnetic Resonance(MR) imaging are the most widely used radiographic techniques in diagnosis, clinical studies and treatment planning. The motive is to discuss the problems encountered in segmentation of CT and MR images, and the relative merits and limitations of methods currently available for segmentation of medical images. With increasing use of CT and MR imaging for diagnosis, treatment planning and clinical studies, it has become almost compulsory to use computers to assist radiological experts in clinical diagnosis, treatment planning. Reliable algorithms are required for the delineation of anatomical structures and other regions of interest.

The techniques available for segmentation of medical images are specific to application, imaging modality and type of body part to be studied. Segmentation based on gray level techniques such as thresholding, and region based are the simplest techniques and find limited applications. However, their performance can be improved by integrating them with artificial intelligence techniques. Techniques based on textural features have excellent results on medical image segmentation. The limitation is that under certain circumstances it becomes difficult to correctly select and label data; has difficulties in segmenting complex structure with variable shape, size, and properties. A variety of different neural network-based algorithms are also available for texture- based segmentation and classification having good accuracy. However, most of these neural network-based algorithms require extensive supervision and training and their performance depends upon the training method and data used in training.

The use of medical image segmentation in a particular mode in which the medical image exist is also described along with the difficulties encountered in each mode. This survey mainly focus on segmentation of Computed Tomography and Magnetic Resonance images. The research focuses on classification of brain tumor medical images. Brain tumor classification consist of three steps-Preprocessing, Feature Extraction and Classification. Based on the features, classifier discriminate
the classes to which it belongs to.

II. LITERATURE SURVEY

A combined technique based on GA and SVM for classification of tumors which was introduced by Sachdeva J et al[3] .428 image dataset were used, out of those, 118 were of Astrocytoma, 59 were of Glioblastoma, and 97,88,66 were Meningioma, Oblastoma, Metastatis respectively. The segmented tumor portion was processed for feature extraction using GLCM in combination with LoG, IBF, RILBP. The reduced feature set was classified using SVM. This method yields an accuracy of 92% approximately. Sachdeva J presented another classifier using ANN and its variants[4]. While using a huge data-set of 428 MR images, claims an accuracy of 85%.

A mix technique consists of a combination of ANN and KNN as a classifier proposed by Badarneh A et al[5]. This work follows mainly three steps, feature extraction using DWT, feature selection using PCA and proposed classifier ANN+KNN. The dataset contains total of 275 MR images of 256 x 256 pixels size. Dataset are categorized into normal (94) and abnormal ,(181) brain MR images. These features are being fed into neural network classifier and KNN classifier separately. The final class label has provided by the majority voting of the two classifiers. The proposed hybrid classifier approach gives good overall results.

A SVM based classification approach was used by Nanthagopal A P[8]. A dual intensity based features are been extracted using DCT and DWT while GA followed by PCA is used for feature selection. Thus on a dataset of 120 images categorized into normal or abnormal MR images a classification accuracy achieved is 96%. While using texture, shape based features with ranking mechanism for feature selection the classifier accuracy fluctuates. Othman B uses PCA for feature selection and PNN for classification[7]. On a dataset of 15 images, accuracy achieved is 80%. When the dataset of images increases, accuracy decreases . It has been noted that SVM gives good results at Intensity based features rather than texture features.

Arakeri M P and Reddy[6] proposed a new indexing based approach called Feature Database Tree (KD- Tree) for faster retrieval of the result in CBIR system. Total 82 images were used, out of those, 42 were of normal images and 40 were malignant ones. This dataset was processed by 10 folds randomly. In each fold, wavelet transformation and modified fuzzy c-means algorithms are used for image segmentation. The segmented tumor portion was processed for feature extraction using GLCM , Shape features followed by feature reduction using PCA . The reduced feature set was classified using ensemble classifier using SVM, ANN, and KNN. The class labels along with feature set are stored in feature database for faster retrieval using KD-Tree indexing. This approach gives the classification accuracy of 97% approximately.

Using the same intensity based features and texture based features, when PNN is applied to a small dataset of 20 images the accuracy achieved is high but as the classifier is tested on a dataset of 120 MR images the accuracy decreases. It has been noted during the experiments that PNN works well for frequency based features set extracted using either DWT or DCT but degrades the performance with texture based features extracted using GLCM. For the classification of tumors, the researchers used a techniques of feature extraction which includes texture and frequency based features extracted using GLCM and DWT which is followed by feature selection techniques like PCA or Genetic algorithm. Main emphasis was given on classification accuracies. For such various researchers have proposed a variants of ANN classifiers. The results shows that the classifiers like ANN, PNN, BPN works well for

frequency based features while classifiers like KNN is good for texture based features. It has also been seen that as the size of dataset varies the corresponding accuracy rate shuffles. With the increase in the dataset of images the corresponding accuracy starts decreasing. Most of the researchers used a dataset of images between 80 to 120 images. There is still a requirement of some new algorithms and techniques which can give better efficiency rate on a larger datasets.

2.1 Thresholding Method

Thresholding based image segmentation[9] aims to partition an input image into pixels of two or more values through comparison of pixel values with the predefined threshold value T individually. Chunyan Jiang et al[10] presents an automatic image segmentation method using thresholding technique. This is based on the assumption that adjacent pixels whose value (grey level, color value, texture, etc) lies within a certain range belong to the same class and thus, good segmentation of images that include only two opposite components can be obtained.

Edge based segmentation[2] is the location of pixels in the image that correspond to the boundaries of the objects seen in the image. It is then assumed that since it is a boundary of a region or an object then it is closed and that the number of objects of interest is equal to the number of boundaries in an image. For precision of the segmentation, the perimeter of the boundaries detected must be approximately equal to that of the object in the input image. Jaskirat K [11] proposed thresholding and edge detection being one of the important aspects of image segmentation comes prior to feature extraction and image recognition system for analyzing images. It helps in extracting the basic shape of an image. Other problems of these techniques emanate from the failure to adjust/calibrate gradient function.

2.2 Pattern Recognition Techniques

Pattern Recognition Techniques is a non-linear modeling tools and can be used to model the inputs and outputs relationships. The extracted information from the training set provides important cues of the structures such as intensity, position and shape, which can be valuable complementary information for the segmentation of test images.

Active appearance models (AAM) are statistical models of the shape of structures. Training samples are used to extract the mean shape, mean appearance and define ranges of shape parameters. Restrictions on shape parameters guarantee the similarity between the segmentation result and the training samples. The segmentation procedure is to find the better positions of the shape points according to the appearance information. Algorithms based on classifiers have been widely applied to segment organs in medical images like cardiac and brain images. If properly modeled, supervised classification algorithms can greatly enhance the segmentation accuracy. However, supervised classification algorithms are sensitive to the initial conditions. To guarantee the correctness of the results, the training set must contain enough samples and the samples should be representative and segmented accurately.

Markovian models of images may help us to make better image restoration, enhancement or segmentation. However, using segmentation methods based on MRF models requires a huge computing power and quite a lot of time. For these reasons, MRF methods are usually used in offline tasks, never in a real-time processing environment. In image processing, MRF modeling has received a great deal of attention in the past decade. This type of modeling, originally introduced in vision by Geman and Geman , has been widely used for edge detection, image restoration, stereo vision, long range motion and image classification.

The SVM[12] is a learning machine for two-group classification problems. The machine conceptually implements the following idea: input vectors are non-linearly mapped to a very high dimension feature space. SVM is a set of supervised learning techniques for solving problems of discrimination, regression and are particularly adapted to data process at very high dimensions. The algorithm of the SVM is described as follows: First specifies a small set of training pixels, such as a small part of an object and a small part of the background, as the clues. Then, fast SVM is applied to train the classifiers based on the training pixels. Finally, the remaining image, which is viewed as the test set, is subdivided into several regions by the classifier.

2.3 Algorithms Based on Deformable Models

Compared with the algorithms of the above two categories, the ones based on deformable models are more flexible and can be used for complex segmentations. According to the representation way of the contour, deformable models can be classified to parametric models and geometric models. A moving equation should be defined to drive the initial contours to the structure boundaries. Therefore, the procedure of these algorithms can be viewed as a modeling of curve evolution.

Parametric deformable models have high computational efficiency and are easy to incorporate with other techniques; Geometric deformable models have the advantage of naturally handling the topological changes. For the medical image segmentation, using parametric model or geometric model depends on the applications. In general, when structures have large shape variety or complicated topology, geometric deformable models are preferred; when the interested structures have open boundaries or the structures are thin or the algorithms need real-time operations, parametric models are preferred. However, deformable models usually contain certain number of parameters. To select proper parameters is critical to the final segmentation results while this is usually a time-consuming task.

III. MEDICAL IMAGES CLASSIFICATION

Medical images classification is a supervised method which is based on probability distribution models that could be parametric or non-parametric such as Euclidean classifier, k-nearest neighbor, minimum distance and decision tree etc. Although, in supervised classification one is provided with a collection of labeled (pre-classified) images and the problem is to label newly encountered, unlabeled images. In general, the prearranged labeled (usually called training set) images are used to do the machine learning of the class description which in turn is used for unknown image . Figure 1 shows the state of the art medical images classification process. Image acquisition involves selection of images ranging from CT, MRI to x-ray etc. Pre-processing is a course of actions that is executed on raw data in order to achieve the best recital for ones datasets. It has significant impact on the performance of classification algorithm. Feature extraction is a process to analyze objects and images to extract the most prominent features that are correspondence of various classes of objects. Therefore, it is worth to state that improving feature extraction process will be likely improving performance of a described classification algorithm. For classification, different applications of data mining techniques are used to predict class membership for data instances.

The development and application of medical image classification has span into different applications of data mining. In view of the fact that classification methods are known as supervised methods because they involve training data that are manually partition and subsequently used as reference for automatically classify new data (test data). Therefore, classification methods can be broadly categorized into three namely: texture classification, neural networks and data mining techniques respectively. Texture classification is an image processing technique which helps to identifies different regions of image by means of texture properties. Neural network are promising alternative to different conventional classification methods. Data mining is one of the realm which uses statistical, machine learning, visualization, other data manipulation and extraction techniques to help simplify data complexity and detect hidden pattern data. In general, the design and selection of classification method needs a careful attention to the following issues: definitions of class/group, Pattern/features representation, feature extraction, feature selection, selection of training and test samples, time complexity and performance evolution.

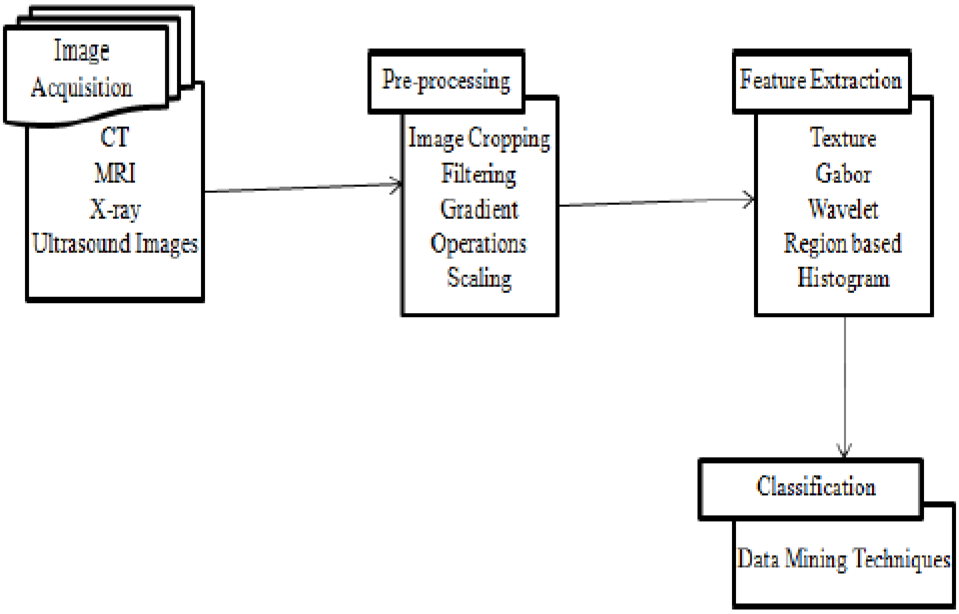

Fig. 1. Medical Images Classification Process

IV. SEGMENTATION TECHNIQUES CLASSIFICATION

Different researchers have done the classification of segmentation techniques in one or another way. At present, from the medical image processing point of view , done the classification of segmentation techniques on the basis of gray level based and textural feature based techniques. Further, consider artificial intelligence as tools to optimize these basic techniques to achieve accurate segmentation results.

4.1 Method Based On Gray Level Features

1) Amplitude segmentation based on histogram features: This includes segmentation of an image based on thresholding of histogram features and gray level thresholding is perhaps the simplest example of this technique. This is particularly suitable for an image with region or object of uniform brightness placed against a background of different gray level. A threshold can be applied to segment the object and background. This gives good results for segmentation of image with bi-modal histogram and fails in the case of an image with multi-modal histogram. Thresholding operation, is very basic and simple, and works well only when the object and background have uniform brightness of distinct gray level values respectively. This simple threshold operation does not work well at segmentation of images with multiple objects each having distinct gray level value varying over a band of values. To overcome this limitation, band thresholding based multiple thresholding operation is applied. For application of thresholding based segmentation technique[9], it is required to apply the correct threshold values in order to achieve proper segmentation results, otherwise results are poor. The major limitation is the selection of proper values of threshold is quite difficult. Performance is affected in presence of artifacts.

2) Edge based segmentation: Edge based segmentation is the most common method based on detection of edges i.e. boundaries which separate distinct regions. Edge detection method is based on marking of discontinuities in gray level, color etc., and often these edges represent boundaries between objects. This method divides an image on the basis of boundaries. Number of edge detecting operators based on gradient (derivative) function are available e.g. Prewitt, Sobel, Roberts (1st derivative type) and Laplacian (2nd derivative type), Canny, Marr-Hilclrath edge detector[2]. Further, in edge based segmentation method, it is required to build the border by combining the detected edges into a edge chain in this process the spurious, or fake edges, weak edges are removed by thresholding operation. The generalized algorithm for edge based segmentation has the following steps:

- Apply the derivative operator to detect edges of the image
- Measure the strength of edges by measuring amplitude of the gradient
- Retain all edge having magnitude greater than threshold T (removal of weak edge)
- Find the position of crack edges; the crack edge is either retained or rejected based on the confidence it receives from it predecessor and successor edges
- Step 3 and 4 are repeated with different values of threshold so as to find out the closed boundaries; segmentation of an image is achieved.

The limitations of edge based method are: Performance is affected by the presence of noise, fake edges and weak edges may be present in the detected edge image which may have a negative influence on segmentation results. Edge detection techniques are required to be used in conjunction with region-based technique for complete segmentation.

3) Region based segmentation: Region based methods are based on the principle of homogeneity- pixels with similar properties are clustered together to form a homogeneous region. The criteria for homogeneity is most of the time gray level of pixels and this criteria can be specified by following conditions:

$$R1UR2U....URi=I$$

where R1, R2,…., Ri are the region in the image I. This is as per the set theory of homogeneity. Region based segmentation is further divided into three types based on the principle of region growing:

- Region merging
- Region splitting
- Split and merge

The algorithm for split and merge follows the following steps:

- Define homogeneity criterion. Break image into four square quadrants.
- If any resultant square is not homogeneous split it further into four quadrants.
- At each level merge the two or more neighboring regions satisfying the condition of homogeneity.
- Continue the split and merge until no further split and merge of region is possible.

Region-growing approaches exploit the important fact that pixels which are close together have similar gray values. Region growing approach is the opposite of the split and merge approach. Region growing methods often give very good segmentations that correspond well to the observed edges. The limitation of region based segmentation is that there are chances of under segmentation and over segmentation of regions in the image. However, this problem can be rectified in two ways- By optimally selecting the criterion for segmentation, for this several algorithm utilizing artificial intelligence techniques have been developed, and, By combining region based approach with edge based approach.

4.2 Method Based on the Textural Features

Textural features[1] of image are important from image segmentation and classification point of view. Different researchers have used these features to achieve image segmentation, classification, and both segmentation as well as classification. The aim of texture based segmentation method is to subdivide the image into region having different texture properties, while in classification the aim is to classify the regions which have already been segmented by one or other method. Texture primitive is a group of pixels representing the simplest or basic sub pattern and follows three main approaches for texture feature extraction based on the type of approach used.

- Statistical approach
- Syntactic or structural approach
- Spectral approach

In case of statistical approach, texture is defined by a set of statistically extracted features represented as vector in multidimensional feature space. The statistical features could be based on first-order, second-order or higher-order statistics of gray level of an image. The feature vector so generated from patterns is assigned to their specific class by probabilistic or deterministic decision algorithm. In case of syntactic approach, texture is defined by texture primitives which are spatially organized according to placement rules to generate complete pattern. In syntactic feature based pattern recognition[16], a formal analogy is drawn between the structural pattern and the syntax of language.

In spectral method, textures are defined by spatial frequencies and evaluated by autocorrelation function of a texture. Some methods available for textural feature extraction and classification based on the above approaches are: co-occurrence matrix method based on statistical description of gray level of an image, gray level run length method, fractal texture description method, syntactic method and Fourier filter method. Comparing the above mentioned three approaches; spectral frequency-based methods are less efficient while statistical methods are particularly useful for random patterns/textures and for complex patterns, syntactic or structural methods give better results. Texture based methods as best suited for segmentation of medical image, when compared to segmentation of medical image using simple gray level based methods.

V. ARTIFICIAL INTELLIGENCE TOOLS FOR SEGMENTATION AND CLASSIFICATION

Automatic segmentation methods have been based on artificial intelligence (AI) based techniques[13]. AI techniques can be classified as supervised and unsupervised. Supervised segmentation requires operator interaction throughout the segmentation process whereas unsupervised methods generally require operator involvement only after segmentation is complete. Unsupervised methods are preferred to ensure a reproducible result; however, operator interaction is still required for error correction in the event of an inadequate result.

5.1 Supervised Methods

In the supervised category, we can place mostly Artificial Neural Network (ANN) based algorithms. ANN[14] is composed of large number of interconnected processing elements (artificial neurons) working in unison to solve specific problems. The main advantages of ANN are: ability to learn adaptively, using training data to solve complex problems. Capability of self-organization; it can create its own organization depending upon the information it receives during learning time capability of performance in real time
because of parallel configuration.

Kailash D. Kharat[15] presented work on brain tumor classification using neural network in which DWT used for feature extraction, dimensionality reduction using PCA and a combination of two classifiers feed forward and back propagation neural networks were used for classification of tumor type with MRI data as input. At times supervised image segmentation and classification methods become very expensive, difficult and even impossible to correctly select and label the training data with its true category. Training is the main requirement of many ANN based algorithms where the classifiers need to be trained before it can be applied to segmentation and classification problem. Further, for different data sets, analysis of different images of different type and format, the whole effort of selecting training data set and training is required to be redone.

5.2 Unsupervised Methods

Most of the unsupervised algorithms are cluster based and not dependent on training and training data. The two commonly used algorithms for clustering are K-mean or Hard C-mean[17] and Fuzzy C-means. K-means algorithm produces results that correspond to hard segmentation while fuzzy C-mean produces soft segmentation which can be converted into hard segmentation by allowing the pixels to have membership of cluster in which they have maximum value of membership coefficients.

In clustering, the aim is to construct decision boundaries based on unlabeled training data. Clustering is the process of finding natural grouping clusters in multidimensional feature space. It is difficult because clusters of different shapes and sizes can occur in multidimensional feature space. A number of functional definitions of clusters have been proposed: Patterns within a cluster are more similar to each other than patterns belonging to different clusters. Image segmentation may be considered a clustering process in which the pixels are classified into the attribute regions based on the texture feature vector calculated around the pixel local neighborhood. Fuzzy clustering is a good method of classifying collection of data point to reside in multiple clusters with different degrees of membership (fuzzy c mean algorithm)[16]. However, the main limitations of fuzzy clustering algorithm are:

- sensitivity to initial partition matrix
- stopping criterion
- solution may get stuck at local minima.

Hence, clustering techniques may not result in optimal solution and there is no best clustering algorithm for a particular application. A number of different algorithms are required to be tried to find the best one.

VI. CONCLUSION

Various techniques on brain tumor medical image classification has been discussed. Recent research in medical image segmentation techniques is presented. After the analysis of different techniques of image segmentation, it is observed that a hybrid solution for image segmentation consists of two or more techniques is being the best approach to solve the problem of image segmentation. Thereby achieving a good classification accuracy. This survey focus on brain tumor classification mechanism. The preprocessing is done to remove noises in the brain image and thereby increasing its quality. Segmentation of medical images is done for obtaining correct target image. Using Texture descriptor, the key features are extracted. After that, the images were trained proposed classifier. Finally, proposed mechanism automatically classifies a query image and overcomes the existing drawbacks of brain image classification.

REFERENCES

[1]. Shwetha Jain,Shubha Mishra, “Artificial Neural Network Approach based on BPN and PNN to classify Brain Cancer”, IJITEE, August 2013.

[2]. Gevers T, Smeulders A W, “Combining region splitting and edge detection through guided Delaunay image subdivision”, IEEE Computer Society, May 2014.

[3]. Sachdeva J, Kumar V, “Segmentation, Feature Extraction and Multi class Brain Tumor Classification”, IEEE ,2011.

[4]. Sachdeva J, Gupta I, “Segmentation, Feature Extraction and Multi class Brain Tumor Classification”, International Journal of Digital Imaging, Springer,2013.

[5]. Badarneh A, Najadat H, “A classifier to detect Tumor Disease in MR Brain Images”, International Conference on Advances in Social Networks Analysis and Mining, IEEE,2012.

[6]. Arakeri M P, Reddy G R M," An intelligent content-based image retrieval system for clinical decision support in brain tumor diagnosis", International Conference on Advances in Social Networks Analysis and Mining, Springer,2013.

[7]. Fauzi M, Othman B, "MRI Brain Classification Using Support Vector Machine", International Conference on Advances in Social Networks Analysis and Mining, IEEE,2011.

[8]. Nanthagopal A P, Rajamony R S, "Classification of benign and malignant brain tumor CT images using wavelet texture parameters and neural network classifier", International Journal of Computer Vision, Springer,2012.

[9]. Ramesh N, Yoo JH, Sethi IK, "Thresholding based on histogram approximation", Proceedings of the International Conference on Vision Image Signal Processing, IEEE,2013.

[10]. Chunyan Jiang,Xinhua Zhang, "Segmentation and Quantification of Brain Tumor", International conference on Virtual Environment,IEEE, July 2000.

[11]. Jaskirat Kaur, Sunil Agrawal, "Comparative Analysis of Thresholding and Edge Detection Segmentation Techniques", International Journal of Computer Applications, February 2012.

[12]. Lahouaoui Lalaoui,Tayeb Mohamadi, "A Comparative Study of Image Region Based Segmentation Algorithms", IJACSA, 2013.

[13]. S N Deepa, B Aruna Devi, "A survey on artificial intelligence approaches for medical image classification", Indian Journal of Science and Technology, November 2011.

[14]. S N Deepa, B Aruna Devi, "Artificial Neural Network Design for Classification of Brain Tumor", IEEE 2012.

[15]. Kailash D Kharat, Pradyumna P Kulkarni, "Brain Tumor Classification using Neural Network based methods", International Journal of Computer Science and Informatics, 2012.

[16]. G Nagalakshmi,S Jyothi, "A Survey on Pattern Recognition using Fuzzy Clustering Approaches", IRJES ,May 2013.

[17]. Pushpa R Suri,Mahak, "Image Segmentation with Modified K-Means Clustering Method", IJRTE, June 2012.